Structural perspective of cooperative transcription factor binding

Morgunova, Ekaterina

2017-12

Morgunova , E & Taipale , J 2017 , ' Structural perspective of cooperative transcription factor binding ' , Current Opinion in Structural Biology , vol. 47 , pp. 1-8 . https://doi.org/10.1016/j.sbi.2017.03.006

http://hdl.handle.net/10138/231316
https://doi.org/10.1016/j.sbi.2017.03.006

Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Structural Biology

Structural perspective of cooperative transcription factor binding

Ekaterina Morgunova[1] and Jussi Taipale[1,2]

In prokaryotes, individual transcription factors (TFs) can recognize long DNA motifs that are alone sufficient to define the genes that they induce or repress. In contrast, in higher organisms that have larger genomes, TFs recognize sequences that are too short to define unique genomic positions. In addition, development of multicellular organisms requires molecular systems that are capable of executing combinatorial logical operations. Co-operative recognition of DNA by multiple TFs allows both definition of unique genomic positions in large genomes, and complex information processing at the level of individual regulatory elements. The TFs can co-operate in multiple different ways, and the precise mechanism used for co-operation determines important features of the regulatory interactions. Here, we present an overview of the structural basis of the different mechanisms by which TFs can cooperate, focusing on insight from recent functional studies and structural analyses of specific TF–TF–DNA complexes.

Addresses
[1] Department of Medical Biochemistry and Biophysics, Karolinska Institutet, SE 171 77 Stockholm, Sweden
[2] Genome-Scale Biology Research Program, P.O. Box 63, FI-00014 University of Helsinki, Finland

Corresponding author: Taipale, Jussi (jussi.taipale@ki.se)

Current Opinion in Structural Biology 2017, **47**:1–8

This review comes from a themed issue on **Protein–nucleic acid interactions**

Edited by **Stephen Cusack** and **Christoph W. Müller**

For a complete overview see the Issue and the Editorial

Available online 24th March 2017

http://dx.doi.org/10.1016/j.sbi.2017.03.006

Introduction

Structural analyses of protein–DNA interactions have provided many insights into the molecular basis of recognition of specific DNA sequences by TFs. The ~1400 mammalian TFs [1] are represented by ~400 unique protein–DNA 3D structures, covering most of the major structural families found in humans. Most analyses have focused on individual monomeric or homodimeric TFs bound to their recognition motifs. Despite the importance of cooperative binding of TFs for the specificity of mammalian gene expression, this area has not been extensively studied at a structural level. Of the 72 heterodimeric TF–TF–DNA complex structures, most represent heterodimers that are formed between proteins prior to DNA binding, and only few structures exist for cooperative complexes whose formation depends on DNA.

In addition to the structural studies of TF–DNA complexes, the interactions between DNA-bound TFs have recently been characterized by high-throughput functional genomic methods such as CAP-SELEX (consecutive affinity-purification systematic evolution of ligands by exponential enrichment) developed in our laboratory [2••]. In CAP-SELEX, two TFs are incubated with random DNA sequences, followed by consecutive affinity purification of the TFs. This results in enrichment of DNA sequences that are bound to both TFs, whose specificities, and the relative orientation and spacing of their motifs can then be determined using next generation sequencing. Initial study of 9400 TF–TF–DNA interactions revealed that DNA-dependent TF–TF interactions are very common, and occur also between different TF structural families. In the study, we found 315 TF–TF interactions represented by 618 heterodimeric motifs. Based on the fraction of pairs tested, we estimated that ~25 000 distinct TF pair specificities can contribute to protein–DNA interactions in cells [2••].

Analyses of TF binding inside eukaryotic cells have also revealed widespread cooperative binding between TFs. Instead of decorating the genome relatively evenly, most TF binding occurs in dense clusters of <1 kb in size that are devoid of nucleosomes. This binding mode was first identified in *Drosophila* cells [3•], and subsequently found to also occur in mammals [4–6]. The clustering occurs largely independently of the functional roles of the TFs in the cells studied [6] and is accompanied by enrichment of TF motifs in the clusters. However, the relatively weak enrichment of motifs alone appears not to be sufficient to explain the widespread binding of TFs to these clusters. These results indicate that cooperativity is an inherent feature of mammalian TF binding to DNA.

The cooperativity can arise by multiple mechanisms. In the simplest case, the TF proteins bind to each other also in the absence of DNA, and bind stronger to DNA together than separately. In a related mechanism, the proteins can interact with each other but with an affinity that is insufficient to form a stable TF dimer in solution. Binding to DNA can facilitate the weak interactions between such TFs, for example by bringing the two

TFs close together, or by altering the conformation of one TF in such a way that its interaction with the other TF is favored.

Cooperativity can also be mediated entirely through DNA, in the absence of direct protein–protein contacts between the TFs. For example, two TFs can prefer DNA that is bent in a similar way, resulting in cooperativity because the energy expended in bending of the DNA by the first protein does not have to be spent again when the second protein binds. Such cooperativity can also be induced more distally, due to the fact that binding of TFs restricts the vibrational freedom of DNA [7•,8]. Finally, much of the cooperativity leading to clustering of TFs in the genome is thought to result from indirect effects. As nucleosome binding generally inhibits binding of TFs, creation of a nucleosome-free region due to binding of initial TF(s) results in formation of naked DNA that is free to bind to other TFs as well.

All of these modes of cooperativity have specific features, and differ from each other in important ways. For example, only some modes can lead to changes in binding specificity of the TF pairs. In addition, some modes act very locally, whereas others can act over a considerable distance. Some modes allow the pairing between the TFs to be highly specific, whereas others are not as dependent on precise spacing between TFs and can support formation of more promiscuous pairings between TFs. In the following paragraphs, we will discuss the features of each of these mechanisms, including structural examples of each case.

Protein-level TF–TF complexes

Many TFs are not able bind DNA as a single monomeric protein. They often contain specific protein–protein interaction domains that bind to each other to form a functional dimer, trimer or tetramer that can be either homomeric or heteromeric in nature. The protein-level interaction allows binding of the complex to DNA due to the increase in avidity and/or cooperative interactions between the subunits. As the proteins are bound together, motifs recognized by the subunits are located close to each other, resulting in binding to sites whose consensus sequences are often palindromic, or less commonly direct repeats.

Some of the largest TF families in humans form protein-level dimers in solution. For example, the basic helix-loop-helix proteins (bHLH), such as MYC, MAX, MAD [9], BMAL-CLOCK [10], and E47-NeuroD [11] dimerize via a long helical interaction domains. Similarly, the basic-leucine zipper family (bZIP) including AP-1 (PDB entry 1JNM, not published), MAFs [12], CREBs [13,14], and CEBPs (PDB entries 2E42, 1GU4, not published) use a leucine zipper domain to form dimers at a protein level (Figure 1a).

Similarly, hepatocyte nuclear factors (HNF1, [16]) p53 [17] and prokaryotic lambda repressor-like proteins [18,19] (Figure 1a) utilize the dimerization through specific dimerization domains. The TFs belonging to the E2F family (E2F1-6 and DP1-4) cannot bind DNA as monomers with high affinity, but show high affinity to DNA as an E2F/DP complex [20] (Figure 1a). The activity of dimeric or multimeric TFs can also be regulated at the level of dimerization. For example, STAT proteins need to be phosphorylated [21,22] to form the respective dimeric forms that bind to DNA.

This mode of interaction is highly specific, allowing specific one to one pairings between proteins (for example NF-Y [23,24], Figure 1a). In several cases, multiple related proteins can also form heterodimers between each other, allowing some combinatorial control. For example the bHLH protein MAX can bind DNA as a homodimer, or as a heterodimer with MYC forming a strong activator, or with MXD1 or MNT, forming repressive complexes [9,25,26].

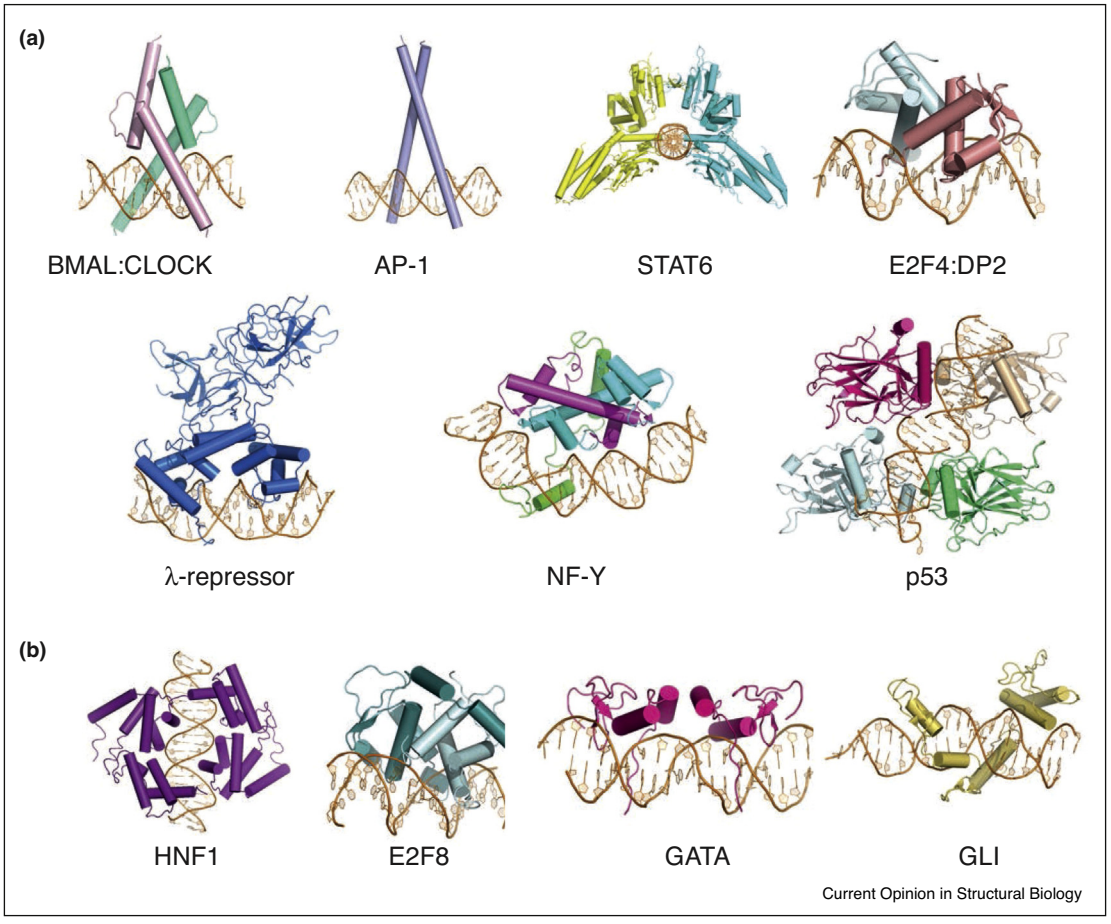
The formation of specific dimers in solution can lead to an absolute requirement of the expression of two specific genes for a particular activity (logical AND gate). This also facilitates robust control of protein activity levels, as the activity of the complex is limited by two concentrations, and overexpression of one subunit alone cannot lead to an increased activity. The protein–protein interaction based mechanism is, however, not very flexible for combinatorial control, as the expression of one highly-expressed A subunit can capture all B subunits and also influence the activity of other A–B pairs. This effect can also be utilized by dominant negative inhibitor proteins such as the ID proteins, which contain only a dimerization domain of a bHLH protein. Expression of these proteins leads to the formation of non-productive complexes, and negatively regulates the activity of many bHLH proteins [27].

In some cases, it appears that the two proteins constituting a heterodimeric TF have fused into a single protein where the two domains are tethered (Figure 1b, HNF1 [16], E2F7 and E2F8 [28], GATA [29], GLI [30] and other C2H2 Zinc finger proteins). This leads to loss of the potential for combinatorial regulation, but facilitates formation of the specific complex and simplifies its regulation.

DNA-facilitated interactions

Some TFs do not interact with each other with appreciable affinity in the absence of DNA, but form highly specific complexes in the presence of DNA. In our CAP-SELEX analyses, we found a large number of such interactions, suggesting that such DNA-facilitated interactions are very common. As two different (asymmetric) TF proteins bound to DNA can be oriented towards each

Figure 1

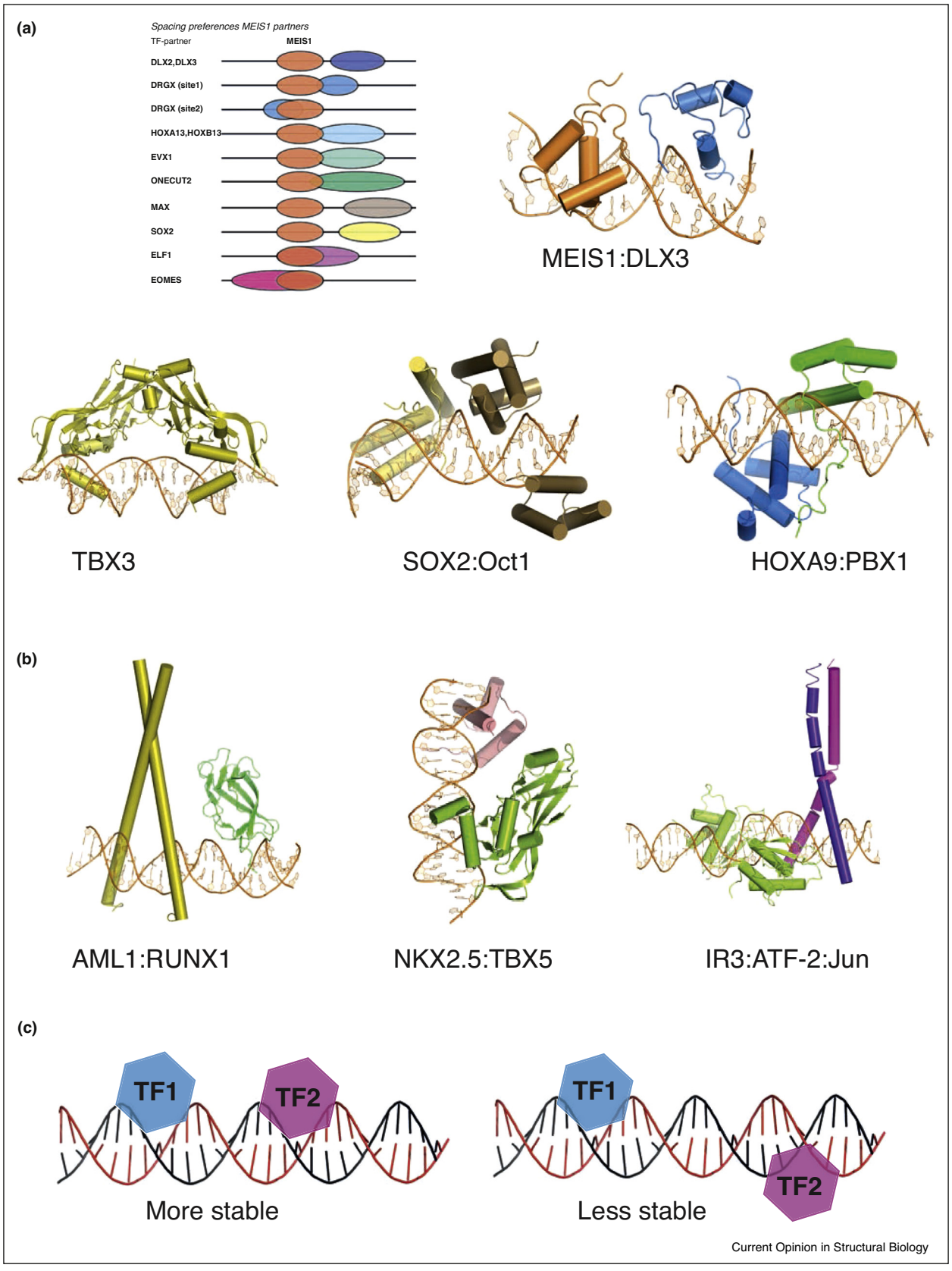


Protein-level TF–TF complexes bound to their specific DNA. **(a)** From left to right in the top line: heterodimer of BMAL(pink):CLOCK(green) (PDB: 4H10); homodimer of AP-1(light blue) (PDB: 1JNM); homodimer of STAT6 (yellow/cyan) (PDB: 4Y5W); heterodimer of E2F4(pink):DP2(light blue) (PDB: 1CF7); on the bottom line: homodimer of lambda-repressor (PDB: 3BDN); trimer of NF-Y (α-subunit in green, β-subunit in cyan and γ-subunit in magenta) (PDB: 4AWL); tetramer of p53 (PDB:2AC0). **(b)** Complexes with two tethered domains: homodimer of HNF1 (PDB: 1IC8); two domains of E2F8 (PDB: 4YO2); two gata-type Zinc −finger domains of GATA (PDB: 3DFV); 3 C2H2-type Zinc-finger domains of GLI (PDB: 2GLI).

other in four different ways (≥>, >≥, ≤>, and ≥<), and their sites can be located at different distances from each other, a large number of potential interacting configurations exist between each TF pair. In addition, as strong cooperative interactions require only formation of one or few hydrogen bonds or van der Waals interactions, the likelihood of such interactions is considerable even in the absence of any selection for the interactions themselves. Using CAP-SELEX, we found that such interactions are indeed commonly observed. For example, DNA-bound TALE homeodomain protein interacted with twelve TFs from six different structural families (Figure 2a).

Structural analysis of one of the MEIS1 interactions, MEIS1:DLX3, revealed a complex interaction involving both DNA and protein [2••]. Insertion of an arginine of DLX3 to the minor groove of DNA lead to immobilization of the peptide backbone in this region in a manner that facilitates hydrogen bond formation between the backbone and an asparagine of MEIS1 [2••]. A simpler example of a DNA-facilitated interaction is illustrated by a complex between the HMG protein Sox2 and the paired domain protein Oct-1 (POU2F1) [2••,31–33]. Binding of the two proteins to DNA places a helix 3 of Sox2 close to the loop between helices 1 and 2 of Oct-1, facilitating interaction between a lysine of Sox2 with main chain glycine of Oct-1 [33]. A combination of a weak protein–protein interaction with a DNA-facilitated mechanism is found in interactions between HOX proteins and TALE homeodomain proteins. Here, a short tryptophan containing peptide motif of the HOX partner associates with a hydrophobic pocket of PBX. The pairing of the proteins is also facilitated by additional protein–protein interactions at the interface that forms

Figure 2



DNA-facilitated interactions. **(a)** In the top line: Positions of MEIS1 partner TFs in relation to the MEIS1 motif (orange, orientation NTGACAN). Panel is adapted from Jolma et al. Nature, 2015, [2••]; cartoon representation of MEIS1(orange): DLX3(blue) complex (PDB: 4XRS); bottom line:

between the two monomers when they bind to DNA [34–36,37•,38–40].

Interestingly, using CAP-SELEX, we found that when two TFs bind to sites that are close to each other, the optimal DNA sequence recognized by them changes from what one would expect from the individual motifs. The structural basis of this effect has not been analyzed, but it primarily affects positions that are recognized indirectly via water-mediated hydrogen bonds or backbone contacts between the TFs and DNA.

The DNA-facilitated interaction mode allows formation of a larger number of TF–TF pairs than direct protein–protein interactions, as its dependence on binding of the individual TFs to their specific sites on DNA prevents the "capture" of TFs by one highly expressed partner. The formation of TF–TF complexes on DNA cannot thus be effectively inhibited by a dominant negative mechanism. As different amino-acids are still involved in the TF–DNA and TF–TF interaction surfaces, this mode still allows evolution of specificity in the TF–TF pairings.

DNA-mediated interactions

The cooperativity between TFs can also be entirely mediated by DNA. In this mechanism, binding of one TF to DNA alters the shape or dynamics of DNA, leading to increased binding affinity of an another TF.

In a simple case, such interactions involve changes in DNA shape that are favored by both of the TFs. In general, hard evidence for such a mechanism is difficult to obtain, as it would require structures of TFs bound separately and together to the same DNA sequence. However, exclusion of other plausible mechanisms by structural studies suggests that DNA-mediated interactions are relatively common. For example, the binary complex of AML1:RUNX1 bound to DNA [41,42], Figure 2b, shows that the proteins do not directly interact with each other, with the shortest distance between amino-acids being above 10 Å. However, the DNA is rather bent between the two bound positions, suggesting that it contributes to the cooperative interactions. Another heterodimeric complex where proteins are relatively far (>6 Å) is the homeobox: T-box complex of NKX2.5:TBX5 [43]. Particular interactions in this complex are shown to be identical to the interactions found in the complexes of the respective individual proteins and DNA. However, mutagenesis study of the residues at the interface between the proteins suggests that some protein–protein interactions occur, potentially indirectly via solvent, or directly between amino-acids due to vibrational movement of the molecules, or alternative conformation of the amino-acids [43].

A classic example of the DNA-mediated interaction mode is also found in the part of the interferon-ß enhanceosome where IR3, ATF-2 and c-Jun DNA-binding domains are bound to DNA (Figure 2b). These proteins do not interact with each other at all directly, but are thought to cooperate through DNA and possibly via co-factors such as CBP [14,44••,45].

Another mechanism by which DNA can mediate cooperativity between TFs is described by Kim *et al.* [7•], who show that TF binding can be stabilized or destabilized by the binding of another TF, and that the effect varies as a function of the distance between the bound sites with a periodicity of ~10 bp, consistent with the helical periodicity of DNA (Figure 2c). A similar periodic pattern was also observed in the CAP-SELEX study. This effect is thought to be caused by quenching of DNA vibrational modes by binding of one protein, resulting in loss of entropy. This entropy cannot be lost again, resulting in cooperativity between TFs that are located up to few helical turns from each other due to the fact that they quench the same vibrational mode(s) [7•,8]. This effect thus has a longer range than the DNA-facilitated and through-DNA modes of cooperativity.

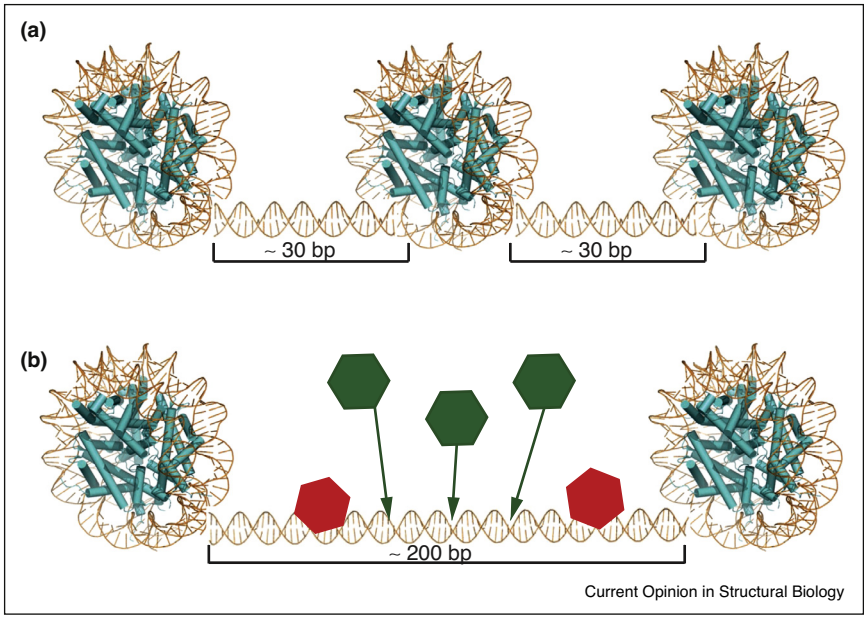
As all DNA-mediated cooperativity modes are caused directly by the interaction of the two TFs with DNA, it is more difficult to achieve a high specificity between TF–TF pairs using this interaction mode. Binding of a TF to DNA thus inevitably leads to changes in the physical properties of DNA, which in turn promote or inhibit binding of all other TFs to sites close to the initially bound site. The specificity of the pairs must thus be determined by the sequence of the DNA itself.

Indirect cooperativity

An alternative mechanism that facilitates binding of multiple TFs close to each other involves competition between nucleosomes and TFs. Whether such an indirect mechanism should be called cooperative depends on definitions, but as the practical consequences are similar to those of directly cooperative binding, we have included discussion of this important mechanism here. The phenomenon of nucleosome-mediated cooperativity has

(Figure 2 Legend Continued) homodimer of TBX3 bound to palindromic site ([15]; PDB: 1H6F) (note that a different dimer may exist without DNA (5BQD)); SOX2(light yellow):Oct1(dark yellow) (PDB: 1O4X); HOXA9(green):PBX1(dark blue) (PDB: 1PUF). **(b)** DNA-mediated interactions. From left to right: AML1(green):RUNX1(yellow) (PDB: 1HJB); NKX2.5(pink):TBX5(green) (PDB: 5FLV); IR3(green):ATF-2(violet):c-JUN(magenta) (PDB: 1T2 K). **(c)** Binding of one TF (TF1) can also regulate binding of other TFs (TF2) at a longer range. This effect is thought to be caused by DNA vibration. During low frequency DNA vibration, major groove widths are correlated along the DNA in such a way that two positions that are a full helical turn apart tend to be both wide at the same time, whereas positions on opposite sides tend to be wide and narrow. Binding of two TFs that favor wide major groove are thus favored on the same side (left), and disfavored on the opposite sides (right). Adapted from (Kim *et al.*, Science, [7•]).

Figure 3



Indirect cooperativity. Schematic presentation of the indirect cooperativity involving competition between nucleosomes (structural representation from PDB: 3AFA) and TFs (colored hexagons). **(a)** Approximately 147 bp of DNA is wound to a nucleosome followed by a linker that is ≥30 bp in length. The linker is accessible, whereas most TFs are not able to bind efficiently DNA wound around the nucleosome. **(b)** When TFs (red) bind and displace the nucleosome in such a way that a region that is shorter than 147 bp is left free, this region cannot accommodate a nucleosome and is thus now accessible for binding to all other TFs (green), leading to apparent cooperativity between the TFs.

been documented by a series of *in vivo* and *in vitro* experiments [2••,46,47,48••,49,50].

Similar indirect mechanism of cooperativity can occur between multiple TFs, and between TFs and non-specific DNA binding proteins such as HMG proteins. Indirect mechanisms can also lead to more complex interactions between DNA-bound proteins. For example, it appears that the architectural protein cohesin that associates with DNA topologically prefers to encircle DNA at nucleosome-free regions [6,51]. However, cohesin does not block binding of TFs, leading to apparent cooperativity between cohesin and TFs.

Indirect cooperativity can act at a relatively long range, as the nucleosome binds to 147 bp of DNA [52]. The effect is relatively non-specific, as generation of nucleosome-free DNA facilitates binding of most DNA-binding proteins. However, TFs can differ in their ability to bind nucleosomal DNA, and thus at least two classes of TFs have been proposed to exist, "pioneer factors" that can displace nucleosomes (reviewed in Ref. [53]), and other TFs that can only effectively bind to free DNA (Figure 3).

Conclusion

The observation that TF binding in cells of higher organisms occurs in dense clusters suggests that most TF binding depends on cooperative interactions. In addition to the classic model of TF–TF cooperativity that depends on protein–protein interactions, recent functional and structural studies have highlighted the important role of DNA in allowing additional modes of cooperation between TFs. DNA facilitates interactions between the TFs themselves, acts to mediate the interactions, and allows indirect cooperativity caused by binding of large architectural proteins such as nucleosomes and cohesin. The cooperativity between TFs allows gene-level processing of combinatorial input information, which is critical for the precise control of gene expression required for development and function of multicellular organisms [54,55]. Structural analysis of the different forms of cooperativity has revealed that mechanistically, the modes have many distinct features (Table 1), and are optimal for different purposes. Together, they allow individual TFs to regulate genes in either simple or complex manner, and individual genes to be regulated either by a single TF, or one or more combinations of TFs. The system thus features both specific and promiscuous mechanisms, allowing both highly accurate regulation, as well as rapid and productive regulatory evolution.

Many recent studies have greatly increased our understanding of the role of TF–TF cooperativity in biological processes. However, multiple key mechanisms remain poorly understood at a molecular level. The central questions include: (1) Why does binding specificity change when TFs bind close to each other? (2) Do

Table 1

Comparison between features of different modes of TF–TF cooperativity

Cooperativity mode	Sequence that primarily determines partner choice	Specificity	Range	Motifs can change
Protein–protein interaction	Protein	Very high	Short	Yes
Tethered complex	Protein	Very high	Short	Yes
DNA-facilitated	Protein and DNA	Moderate	Short	Yes
DNA-mediated	DNA	Low	Short	Yes
DNA-allostery	DNA	Low	Tens of bp	No
Nucleosome-facilitated	DNA	Low	~150 bp	No

non-specific and weakly specific DNA binding proteins such as the HMG factors and AT-hook containing proteins commonly cooperate with TFs? (3) Does TF–TF cooperativity commonly involve allosteric interactions with non-DNA binding cofactors? (4) Do large proteins or protein complexes involved in transcription, such as chromatin modifying enzymes or the mediator influence TF cooperativity? (5) What is the role of DNA looping as found in the case of the lambda repressor [56] in TF–TF cooperativity? (6) Are regulatory elements occupied by 'TF chains' like those found at the enhanceosome, or are they bound by more isolated TFs and cooperative pairs? And finally, (7) What is the molecular basis of enhancer-promoter interactions? Taken together, much work remains in understanding the structural basis of cooperative interactions that are central for both development and disease. The recent advances in functional genomics methods that allow identification of high affinity complexes, and preferential spacings and orientations between specific DNA elements, together with the ability of new electron microscopes to analyze large molecular complexes in atomic detail have made addressing many of these questions feasible in the near future.

Acknowledgements

This work was supported by the Center for Innovative Medicine at Karolinska Institutet, Cancerfonden, the Knut and Alice Wallenberg Foundation and the Swedish Research Council. We thank Drs. Minna Taipale and Otto Kauko for critical review of the manuscript.

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Vaquerizas JM, Kummerfeld SK, Teichmann SA, Luscombe NM: **A census of human transcription factors: function, expression and evolution**. *Nat Rev Genet* 2009, **10**:252-263.

2. •• Jolma A, Yin Y, Nitta KR, Dave K, Popov A, Taipale M, Enge M, Kivioja T, Morgunova E, Taipale J: **DNA-dependent formation of transcription factor pairs alters their binding specificity**. *Nature* 2015, **527**:384-388.

This paper describes a systematic analysis of TF–TF interactions in the presence of DNA, and structural analyses of some DNA-mediated and DNA-facilitated interactions.

3. • Moorman C, Sun LV, Wang J, de Wit E, Talhout W, Ward LD, Greil F, Lu XJ, White KP, Bussemaker HJ *et al.*: **Hotspots of transcription factor colocalization in the genome of *Drosophila melanogaster***. *Proc Natl Acad Sci U S A* 2006, **103**:12027-12032.

This paper describes clustering of TFs close to each other in Drosophila.

4. Garber M, Yosef N, Goren A, Raychowdhury R, Thielke A, Guttman M, Robinson J, Minie B, Chevrier N, Itzhaki Z *et al.*: **A high-throughput chromatin immunoprecipitation approach reveals principles of dynamic gene regulation in mammals**. *Mol Cell* 2012, **47**:810-822.

5. Gerstein MB, Kundaje A, Hariharan M, Landt SG, Yan KK, Cheng C, Mu XJ, Khurana E, Rozowsky J, Alexander R *et al.*: **Architecture of the human regulatory network derived from ENCODE data**. *Nature* 2012, **489**:91-100.

6. Yan J, Enge M, Whitington T, Dave K, Liu J, Sur I, Schmierer B, Jolma A, Kivioja T, Taipale M *et al.*: **Transcription factor binding in human cells occurs in dense clusters formed around cohesin anchor sites**. *Cell* 2013, **154**:801-813.

7. • Kim S, Brostromer E, Xing D, Jin J, Chong S, Ge H, Wang S, Gu C, Yang L, Gao YQ *et al.*: **Probing allostery through DNA**. *Science* 2013, **339**:816-819.

A proposed method of intermediate range TF–TF cooperativity that is caused by DNA vibration.

8. Crothers DM: **Biophysics. Fine tuning gene regulation**. *Science* 2013, **339**:766-767.

9. Nair SK, Burley SK: **Structural aspects of interactions within the Myc/Max/Mad network**. *Curr Topics Microbiol Immunol* 2006, **2**:123-143.

10. Wang Z, Wu Y, Li L, Su XD: **Intermolecular recognition revealed by the complex structure of human CLOCK-BMAL1 basic helix-loop-helix domains with E-box DNA**. *Cell Res* 2013, **23**:213-224.

11. Longo A, Guanga GP, Rose RB: **Crystal structure of E47-NeuroD1/beta2 bHLH domain-DNA complex: heterodimer selectivity and DNA recognition**. *Biochemistry* 2008, **47**:218-229.

12. Katsuoka F, Yamamoto M: **Small Maf proteins (MafF, MafG, MafK): history, structure and function**. *Gene* 2016, **586**:197-205.

13. Schumacher MA, Goodman RH, Brennan RG: **The structure of a CREB bZIP.somatostatin CRE complex reveals the basis for selective dimerization and divalent cation-enhanced DNA binding**. *J Biol Chem* 2000, **275**:35242-35247.

14. Panne D, Maniatis T, Harrison SC: **Crystal structure of ATF-2/c-Jun and IRF-3 bound to the interferon-beta enhancer**. *EMBO J* 2004, **23**:4384-4393.

15. Coll M, Seidman JG, Muller CW: **Structure of the DNA-bound T-box domain of human TBX3, a transcription factor responsible for ulnar-mammary syndrome**. *Structure* 2002, **10**:343-356.

16. Chi YI, Frantz JD, Oh BC, Hansen L, Dhe-Paganon S, Shoelson SE: **Diabetes mutations delineate an atypical POU domain in HNF-1alpha**. *Mol Cell* 2002, **10**:1129-1137.

17. Kitayner M, Rozenberg H, Kessler N, Rabinovich D, Shaulov L, Haran TE, Shakked Z: **Structural basis of DNA recognition by p53 tetramers**. *Mol Cell* 2006, **22**:741-753.

18. Aggarwal AK, Rodgers DW, Drottar M, Ptashne M, Harrison SC: **Recognition of a DNA operator by the repressor of phage 434: a view at high resolution**. *Science* 1988, **242**:899-907.

19. Stayrook S, Jaru-Ampornpan P, Ni J, Hochschild A, Lewis M: **Crystal structure of the lambda repressor and a model for**

pairwise cooperative operator binding. *Nature* 2008, **452**:1022-1025.

20. Zheng N, Fraenkel E, Pabo CO, Pavletich NP: **Structural basis of DNA recognition by the heterodimeric cell cycle transcription factor E2F-DP**. *Genes Dev* 1999, **13**:666-674.

21. Mayr B, Montminy M: **Transcriptional regulation by the phosphorylation-dependent factor CREB**. *Nat Rev Mol Cell Biol* 2001, **2**:599-609.

22. Li J, Rodriguez JP, Niu F, Pu M, Wang J, Hung LW, Shao Q, Zhu Y, Ding W, Liu Y *et al.*: **Structural basis for DNA recognition by STAT6**. *Proc Natl Acad Sci U S A* 2016, **113**:13015-13020.

23. Nardini M, Gnesutta N, Donati G, Gatta R, Forni C, Fossati A, Vonrhein C, Moras D, Romier C, Bolognesi M *et al.*: **Sequence-specific transcription factor NF-Y displays histone-like DNA binding and H2B-like ubiquitination**. *Cell* 2013, **152**:132-143.

24. Nardone V, Chaves-Sanjuan A, Nardini M: **Structural determinants for NF-Y/DNA interaction at the CCAAT box**. *Biochimi Biophys Acta* 2016. Epub ahead of print http://www.sciencedirect.com/science/article/pii/S1874939916301936.

25. Link JM, Hurlin PJ: **The activities of MYC, MNT and the MAX-interactome in lymphocyte proliferation and oncogenesis**. *Biochim Biophys Acta* 2015, **1849**:554-562.

26. Grandori C, Cowley SM, James LP, Eisenman RN: **The Myc/Max/Mad network and the transcriptional control of cell behavior**. *Annu Rev Cell Dev Biol* 2000, **16**:653-699.

27. Benezra R, Davis RL, Lockshon D, Turner DL, Weintraub H: **The protein Id: a negative regulator of helix-loop-helix DNA binding proteins**. *Cell* 1990, **61**:49-59.

28. Morgunova E, Yin Y, Jolma A, Dave K, Schmierer B, Popov A, Eremina N, Nilsson L, Taipale J: **Structural insights into the DNA-binding specificity of E2F family transcription factors**. *Nat Commun* 2015, **6**:10050.

29. Pavletich NP, Pabo CO: **Crystal structure of a five-finger GLI-DNA complex: new perspectives on zinc fingers**. *Science* 1993, **261**:1701-1707.

30. Bates DL, Chen Y, Kim G, Guo L, Chen L: **Crystal structures of multiple GATA zinc fingers bound to DNA reveal new insights into DNA recognition and self-association by GATA**. *J Mol Biol* 2008, **381**:1292-1306.

31. Klemm JD, Pabo CO: **Oct-1 POU domain-DNA interactions: cooperative binding of isolated subdomains and effects of covalent linkage**. *Genes Dev* 1996, **10**:27-36.

32. Remenyi A, Lins K, Nissen LJ, Reinbold R, Scholer HR, Wilmanns M: **Crystal structure of a POU/HMG/DNA ternary complex suggests differential assembly of Oct4 and Sox2 on two enhancers**. *Genes Dev* 2003, **17**:2048-2059.

33. Williams DC Jr, Cai M, Clore GM: **Molecular basis for synergistic transcriptional activation by Oct1 and Sox2 revealed from the solution structure of the 42-kDa Oct1 and Sox2 revealed from the solution structure of the 42-kDa Oct1. Sox2. Hoxb1-DNA ternary transcription factor complex**. *J Biol Chem* 2004, **279**:1449-1457.

34. LaRonde-LeBlanc NA, Wolberger C: **Structure of HoxA9 and Pbx1 bound to DNA: Hox hexapeptide and DNA recognition anterior to posterior**. *Genes Dev* 2003, **17**:2060-2072.

35. Piper DE, Batchelor AH, Chang CP, Cleary ML, Wolberger C: **Structure of a HoxB1-Pbx1 heterodimer bound to DNA: role of the hexapeptide and a fourth homeodomain helix in complex formation**. *Cell* 1999, **96**:587-597.

36. Merabet S, Saadaoui M, Sambrani N, Hudry B, Pradel J, Affolter M, Graba Y: **A unique Extradenticle recruitment mode in the Drosophila Hox protein Ultrabithorax**. *Proc Natl Acad Sci U S A* 2007, **104**:16946-16951.

37. • Mann RS, Affolter M: **Hox proteins meet more partners**. *Curr Opin Genet Dev* 1998, **8**:423-429.
An extensive review on the molecular basis of homeodomain-DNA interactions, including the role of cofactors and cooperativity in DNA binding.

38. Merabet S, Mann RS: **To be specific or not: the critical relationship between Hox and TALE proteins**. *Trends Genet* 2016, **32**:334-347.

39. Bobola N, Merabet S: **Homeodomain proteins in action: similar DNA binding preferences, highly variable connectivity**. *Curr Opin Geneti Dev* 2016, **43**:1-8.

40. Sanchez M, Jennings PA, Murre C: **Conformational changes induced in Hoxb-8/Pbx-1 heterodimers in solution and upon interaction with specific DNA**. *Mol Cell Biol* 1997, **17**:5369-5376.

41. Tahirov TH, Inoue-Bungo T, Morii H, Fujikawa A, Sasaki M, Kimura K, Shiina M, Sato K, Kumasaka T, Yamamoto M *et al.*: **Structural analyses of DNA recognition by the AML1/Runx-1 Runt domain and its allosteric control by CBFbeta**. *Cell* 2001, **104**:755-767.

42. Tahirov TH, Inoue-Bungo T, Sasaki M, Shiina M, Kimura K, Sato K, Kumasaka T, Yamamoto M, Kamiya N, Ogata K: **Crystallization and preliminary X-ray analyses of quaternary, ternary and binary protein–DNA complexes with involvement of AML1/Runx-1/CBFalpha Runt domain, CBFbeta and the C/EBPbeta bZip region**. *Acta Crystallogr Sect D Biol Crystallogr* 2001, **57**:850-853.

43. Luna-Zurita L, Stirnimann CU, Glatt S, Kaynak BL, Thomas S, Baudin F, Samee MA, He D, Small EM, Mileikovsky M *et al.*: **Complex interdependence regulates heterotypic transcription factor distribution and coordinates cardiogenesis**. *Cell* 2016, **164**:999-1014.

44. •• Panne D: **The enhanceosome**. *Curr Opin Struct Biol* 2008, **18**:236-242.
A strucural analysis of the interferon beta enhanceosome, showing many different modes of TF–TF cooperativity in a single regulatory element.

45. Panne D, Maniatis T, Harrison SC: **An atomic model of the interferon-beta enhanceosome**. *Cell* 2007, **129**:1111-1123.

46. Adams CC, Workman JL: **Binding of disparate transcriptional activators to nucleosomal DNA is inherently cooperative**. *Mol Cell Biol* 1995, **15**:1405-1421.

47. Miller JA, Widom J: **Collaborative competition mechanism for gene activation in vivo**. *Mol Cell Biol* 2003, **23**:1623-1632.

48. •• Mirny LA: **Nucleosome-mediated cooperativity between transcription factors**. *Proc Natl Acad Sci U S A* 2010, **107**:22534-22539.
The mechanisms behind indirect cooperativity caused by competition between TFs and nucleosomes are modeled here.

49. Pettersson M, Schaffner W: **Synergistic activation of transcription by multiple binding sites for NF-kappa B even in absence of co-operative factor binding to DNA**. *J Mol Biol* 1990, **214**:373-380.

50. Vashee S, Melcher K, Ding WV, Johnston SA, Kodadek T: **Evidence for two modes of cooperative DNA binding in vivo that do not involve direct protein–protein interactions**. *Curr Biol* 1998, **8**:452-458.

51. Lopez-Serra L, Kelly G, Patel H, Stewart A, Uhlmann F: **The Scc2-Scc4 complex acts in sister chromatid cohesion and transcriptional regulation by maintaining nucleosome-free regions**. *Nat Genet* 2014, **46**:1147-1151.

52. Luger K, Mader AW, Richmond RK, Sargent DF, Richmond TJ: **Crystal structure of the nucleosome core particle at 2.8 A resolution**. *Nature* 1997, **389**:251-260.

53. Zaret KS, Carroll JS: **Pioneer transcription factors: establishing competence for gene expression**. *Genes Dev* 2011, **25**:2227-2241.

54. Spitz F, Furlong EE: **Transcription factors: from enhancer binding to developmental control**. *Nat Rev Genet* 2012, **13**:613-626.

55. Lagha M, Bothma JP, Levine M: **Mechanisms of transcriptional precision in animal development**. *Trends Genet* 2012, **28**:409-416.

56. Ptashne M: **Gene regulation by proteins acting nearby and at a distance**. *Nature* 1986, **322**:697-701.